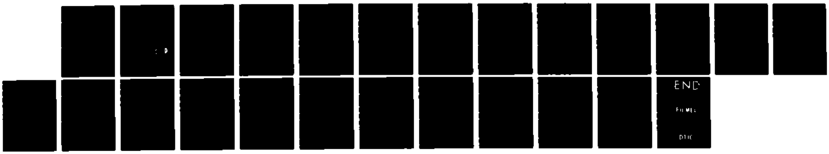
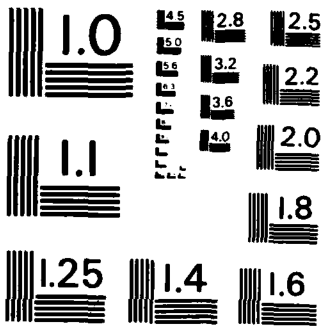
AD-A161 262 A USEFUL CLASS OF MULTISTAGE TESTS(U) ILLINOIS UNIV AT URBANA COORDINATED SCIENCE LAB H V POOR ET AL. DEC 81 R-923 NSF-ECS79-16453 1/1

UNCLASSIFIED F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(12)

REPORT R-923 DECEMBER, 1981 UILU-ENG 81-2254

A USEFUL CLASS OF MULTISTAGE TESTS

H. V. POOR
S. TANTARATANA

APPROVED FOR PUBLIC RELEASE. DISTRIBUTION UNLIMITED.

UN N

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)

A USEFUL CLASS OF MULTISTAGE TESTS

5. TYPE OF REPORT & PERIOD COVERED

Technical Report

6. PERFORMING ORG. REPORT NUMBER

R-923; UILU-ENG-81-2254

7. AUTHOR(s)

H. V. Poor and S. Tantaratana

8. CONTRACT OR GRANT NUMBER(s)

ECS-79-16453

9. PERFORMING ORGANIZATION NAME AND ADDRESS

Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS

National Science Foundation

12. REPORT DATE

December 1981

13. NUMBER OF PAGES

21

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)

Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Multistage tests; asymptotic efficiencies; nonparametric tests

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report considers a class of hypothesis tests in which the (null) hypothesis is rejected only after being rejected on the basis of each of two samples. The motivation for these procedures is that their average sample sizes under the null hypothesis will be very near that of a fixed-sample-size test using only the first sample, while their error-probability performance should be superior. The validity of these properties is verified numerically for the particular problem of location testing with normal errors
(over)

DD FORM 1 JAN 73 1473

20. (continued)

by considering the power functions and average sample sizes of linear versions of these tests. Also, the Pitman asymptotic efficiencies of the proposed tests are compared to those of fixed-sample-size tests and sequential probability ratio tests, indicating that the proposed tests are intermediate to these two alternatives. However, the complexity and other properties of the proposed tests are comparable to those of the fixed-sample-size test, making their use desirable for many applications.^ Nonparametric versions of these tests, based on the sign-test statistic, are also proposed for the location-testing problem, and these are compared to the corresponding fixed-sample-size sign test with essentially the same conclusions being drawn as in the case of the linear tests. Finally, the efficiency of a k-stage version of one of the proposed tests is considered with the conclusion that the behavior of the two-stage version (whether it be good or bad) is amplified by adding stages up to a point beyond which the effects of adding stages diminishes.

UILU-ENG-81-2254

A USEFUL CLASS OF MULTISTAGE TESTS

by

H. Vincent Poor and Sawasd Tantaratana

This work was supported in part by the National Science Foundation under Contract ECS-79-16453.

1. Introduction

In this paper we consider the problem of testing an hypothesis H_0 versus an alternative H_1 for situations in which a large number of tests are to be performed and in which H_1 occurs only rarely. Such situations arise in many applications. For example, in search radar, the radar examines a large number of resolution cells on each scan of a search area, but it is expected that targets will be present in only a small fraction of the resolution cells. Similar situations arise in applications such as medical testing and quality control.

The tests we consider here are described as follows: For each test, a sample is taken and a standard procedure is applied to decide whether or not H_0 can be accepted. If H_0 is accepted, then the test ends. If H_0 is not accepted, then a second sample is taken from the same population and a second test is performed to double-check the result. Thus, H_0 is rejected only if it is rejected on the basis of both samples. The motivation for this type of testing is that, for small probability of Type I error on the first check, the average sample size under H_0 will be very nearly that of the first sample, whereas the overall performance should be superior to a test using only the first sample. The trade-off, of course, is that the average sample size under H_1 may be larger than that required for a comparable fixed-sample-size test; however, since H_1 is assumed to occur only rarely, the overall average sample size should be smaller than that of a comparable fixed-sample-size test.

In this paper, we consider two such double-check procedures — Procedure 1 in which the double-check is performed only on the second sample, and Procedure 2 in which the double-check is performed on the first and second samples combined. To investigate the properties of these tests, we consider

primarily the specific problem of location testing with i.i.d. normal errors. We compare the power functions and average sample sizes of these procedures to those of fixed-sample-size tests of the same significance level and with various sample sizes. We also consider the Pitman asymptotic efficiency of these two procedures relative to both fixed-sample-size tests and sequential tests. Nonparametric versions of the new procedures are also proposed and an analysis of their behavior is included. Finally, a generalization is considered in which H_0 is rejected only after being rejected on k samples where k is a positive integer.

2. General Description and Error-Probability Performance

Suppose we have two samples $X_i = \theta + \epsilon_i$, $i = 1,2,\ldots,n_1$ and $X_i = \theta + \epsilon_i$, $i = n_1 + 1,\ldots,n$, where $\epsilon_1,\ldots,\epsilon_n$ is an i.i.d. sequence of $\mathcal{N}(0,1)$ errors and where $\theta \geq 0$. Consider the two tests for the hypothesis $H_0:\theta = 0$ versus the alternative $H_1:\theta > 0$,

$$\varphi_{DC1}(\underset{\sim}{x}) = \begin{cases} 1; \text{ if } \sum_{i=1}^{n_1} x_i \geq \tau_1 \text{ and } \sum_{i=n_1+1}^{n} x_i \geq \tau_2 \\ 0; \text{ otherwise} \end{cases} \tag{1}$$

and

$$\varphi_{DC2}(\underset{\sim}{x}) = \begin{cases} 1; \text{ if } \sum_{i=1}^{n_1} x_i \geq \tau_1 \text{ and } \sum_{i=1}^{n} x_i \geq \tau_2' \\ 0; \text{ otherwise} \end{cases} \tag{2}$$

where $\varphi_{DC1}(\underset{\sim}{x})$ [resp. $\varphi_{DC2}(\underset{\sim}{x})$] is the probability with which we accept H_1 given that $(X_1,\ldots,X_n) = (x_1,\ldots,x_n) = \underset{\sim}{x}$, and τ_1 and τ_2 [resp. τ_2'] are chosen to give desired probability of Type I error. Note that size α can be achieved by choosing τ_1 to yield size $\alpha^* \geq \alpha$ on the first check (i.e., $\tau_1 = \sqrt{n_1}\,\Phi^{-1}(1-\alpha^*)$ where Φ denotes the unit normal distribution function) and then choosing τ_2 or τ_2' to give overall size α. For (1), the second threshold is thus given by $\tau_2 = \sqrt{n-n_1}\,\Phi^{-1}(1-\alpha/\alpha^*)$, and for (2) the second threshold is given by $\tau_2' = \sqrt{n}\,b$ where b is the solution to the equation

$$\alpha-\alpha^* = F(\Phi^{-1}(1-\alpha^*),b) - \Phi(b), \tag{3}$$

where F denotes the joint unit normal distribution function with correlation coefficient $(1+K)^{-\frac{1}{2}}$ where $K = (n-n_1)/n_1$.

Note that the average sample size under H_0 of either (1) or (2) is given by $(1+\alpha^*K)n_1$, and the average sample size under H_1 of either test is given by $(1+\beta^*K)n_1$, where $\beta^* = 1-\Phi(\Phi^{-1}(1-\alpha^*)-\sqrt{n_1}\,\theta)$. Also, the

power functions (i.e., power versus θ) of φ_{DC1} and φ_{DC2} are easily computed using standard techniques. The behavior of φ_{DC1} and φ_{DC2} was compared to that of the fixed-sample-size test

$$\varphi_{FSS}(\underline{x}) = \begin{cases} 1; \text{ if } \sum_{i=1}^{M} x_i \geq \sqrt{M}\,\Phi^{-1}(1-\alpha) \\ 0; \text{ otherwise} \end{cases} \qquad (4)$$

for a variety of choices of the test parameters α, α^*, K, and M and for $n = 20$. Table 1 contains data for the particular cases $\alpha = 0.01$, $n_1 = 8$, $\alpha^* = 0.5$ and 0.2 and M = 10, 15, and 20. These values of n_1 and α^* are used because they appear to be nearly optimum for φ_{DC2} with this choice of α and n, and the behavior exhibited in Table 1 is typical of the general optimum behavior of φ_{DC1} and φ_{DC2}. Note from Table 1 that, for the choice $\alpha^* = 0.5$, φ_{DC1} and φ_{DC2} have power functions very near those of φ_{FSS} with $M = 15$ and 20, respectively. In this case, φ_{DC2} is clearly superior to the comparable fixed-sample-size test, whereas φ_{DC1} is not. Alternately, for the choice $\alpha^* = 0.2$, φ_{DC1} has a power function which is greater than that of the $M = 15$ version of φ_{FSS} while requiring only 10.4 samples on the average under H_0. Note that the $M = 10$ version of φ_{FSS} is not comparable to φ_{DC1} in this case. Note also that, with $\alpha^* = 0.2$, φ_{DC2} still compares favorably with the $M = 20$ version of φ_{FSS} in terms of power while rataining the small average sample size under H_0. It is noteworthy that the average sample sizes of the double-check tests can in no case exceed n under either hypothesis.

3. Asymptotic Efficiencies of the Double-Check Tests

Another comparison of interest is to consider the Pitman asymptotic efficiencies (as defined, for example, in Noether (1955)) of the tests (1) and (2) relative to (4). In particular it is straightforward to show that, for fixed α and power β, under H_0 we have

$$ARE^0_{DC,FSS} = \frac{[\Phi^{-1}(1-\alpha) - \Phi^{-1}(1-\beta)]^2}{(1+\alpha^*K)R^2} \tag{5}$$

where, for φ_{DC1}, the parameter R is the solution to the equation

$$\beta = \Phi(R - \Phi^{-1}(1-\alpha^*))\Phi(\sqrt{K}R - \Phi^{-1}(1-\alpha/\alpha^*)), \tag{6}$$

and, for φ_{DC2}, R solves

$$\beta = \Phi(R - \Phi^{-1}(1-\alpha^*)) - \Phi(b - \sqrt{1+K}R) + F(\Phi^{-1}(1-\alpha^*) - R, b - \sqrt{1+K}R), \tag{7}$$

where b and F are as in (3). Similarly, under H_1 we have

$$ARE^1_{DC,FSS} = \frac{(1+\alpha^*K)}{(1+\beta^*K)} ARE^0_{DC,FSS} \tag{8}$$

where β^* is as defined above.

Values of $ARE^j_{DC,FSS}$ for $j=0$ and 1 are given in Table 2 for a variety of values of α^*, K, α, and β. Note that φ_{DC2} is uniformly superior to φ_{FSS} under H_0 for the ranges of parameters considered. Furthermore, in each case, values of α^* and K can be chosen so that φ_{DC2} is nearly as efficient as φ_{FSS} under H_1 as well. As expected, φ_{DC1} is slightly less efficient than φ_{DC2}, but α^* and K can be chosen to yield efficiency under H_0 higher than that of φ_{FSS} in each case considered. The conditions of Table 2(c) (i.e., $\alpha = 10^{-6}$, $\beta = 0.95$) appear to be most favorable for

performance under H_0 of either double-check test. These general conditions (i.e., very small α and moderate $(1-\beta)$) are the most prevalent for many testing problems (such as that arising in search radar). The antipodal conditions (very small $(1-\beta)$ and moderate α) are less favorable for performance of the double-check tests. However, these latter conditions carry the implication that Type II errors are more significant than Type I errors, and thus that the double-check should be performed only when rejecting H_1 rather than when rejecting H_0. This would result in the performance tabulated in Table 2(c) with the roles of H_0 and H_1 reversed.

Of course, the optimum multistage test (in terms of average sample size) is the Wald sequential probability ratio test (SPRT). The asymptotic efficiency of SPRT relative to φ_{FSS} has been considered by Paulson (1947) and Bechhofer (1960), and by combining their results with (5) and (8) the asymptotic efficiencies of φ_{DC1} and φ_{DC2} relative to the SPRT can be computed straightforwardly. Typical values of the asymptotic efficiency of φ_{DC2} (with $\alpha^* = 3\sqrt{\alpha}$ and $K = 2$) relative to the SPRT are given in Table 3. Note that these values range from approximately 36% to approximately 65% under H_0 and from approximately 11% to approximately 36% under H_1. Thus, as is expected, the SPRT is superior in performance to φ_{DC2}. However, the double-check tests are still preferable to the SPRT for many applications for several reasons. First, the double-check test is much simplier to implement since it requires at most two comparisons. Further, the thresholds of the double-check tests can be set without knowledge of the true value of θ under H_1. This is not true of the SPRT. Moreover, the SPRT can be less efficient than even the fixed-sample-size test if an incorrect value of θ is assumed (Wald (1947)). Finally, the maximum value of the sample size is finite for the double-check test, whereas the sample size of the

SPRT can assume any positive integral value (although this disadvantage of the SPRT can, of course, be eliminated by truncation).

4. A Nonparametric Version of φ_{DC}

Nonparametric versions of (1) and (2) are easily constructed by replacing the linear statistics with (H_0) distribution-free statistics. For example, if X_i is replaced by $\mathrm{sgn}(X_i)$ and randomization is introduced on the threshold boundaries, then the tests of (1) and (2) become nonparametric for the hypothesis $P(X_i < 0) = \frac{1}{2}$. The power functions (for the alternative $P(X_i \geq 0) = p > \frac{1}{2}$) of these particular nonparametric versions of (1) and (2) are compared to that of the fixed-sample-size sign test (i.e., (4) with X_i replaced by $\mathrm{sgn}(X_i)$ and with randomization on the threshold boundary) in Table 4. Note that the Pitman asymptotic relative efficiencies (for location testing) between these modified tests will be the same as those for the linear tests within mild regularity conditions on the distribution of the ϵ_i (such as those given in Noether (1955)). It is also noteworthy that, again within regularity, the asymptotic location-testing efficiencies of these double-check sign tests relative to the linear test of (4) are $4f^2(0)\mathrm{Var}(\epsilon_i)$ times the value computed from (5) and (8), where f is the probability density of ϵ_i. For the case of normal errors we have $4f^2(0)\mathrm{Var}(\epsilon_i) = 2/\pi$, and thus, whenever a value from Table 2 exceeds $\pi/2 \cong 1{\cdot}57$, the (nonparametric) double-check sign test is more efficient than the Neyman-Pearson test for normal errors.

5. k-Stage Tests

The above analysis of the double-check tests of (1) and (2) can be extended straightforwardly to tests which reject H_0 only after k samples. For example, consider the following test based on kn_1 i.i.d. observations:

$$\varphi_{kC}(\underline{x}) = \begin{cases} 1; \text{ if } \min\{\sum_{i=1}^{n_1} x_i, \sum_{i=n_1+1}^{2n_1} x_i, \ldots, \sum_{i=(k-1)n_1+1}^{kn_1} x_i\} \geq \sqrt{n_1}\,\Phi^{-1}(1-\alpha^{1/k}) \\ 0; \text{ otherwise} \end{cases} \tag{9}$$

Note that the case $k=2$ is φ_{DC1} with $\alpha^* = \sqrt{\alpha}$ and $K=1$. In general this test distributes the Type I error uniformly over the k samples. The asymptotic efficiency under H_0 of (9) relative to (4) is given straightforwardly by

$$ARE^0_{kC,FSS} = \frac{(1-\alpha^{1/k})[\Phi^{-1}(1-\alpha) - \Phi^{-1}(1-\beta)]^2}{(1-\alpha)[\Phi^{-1}(1-\alpha^{1/k}) - \Phi^{-1}(1-\beta^{1/k})]^2}, \tag{10}$$

and under H_1 we have

$$ARE^1_{kC,FSS} = \frac{(1-\beta^{1/k})(1-\alpha)}{(1-\beta)(1-\alpha^{1/k})} \; ARE^0_{kC,FSS} \, . \tag{11}$$

Table 5 gives values of the quantities of (10) and (11) for several values of α and β and for values of k from 2 to 10. Note that the addition of stages improves performance under H_0 to a point (significantly in some cases), but that there is a diminishing return and even decreased performance associated with larger numbers of stages. Also note that, as one might expect, the performance under H_1 degrades with increasing k. As a general rule, it appears that the addition of more stages is helpful in those cases

in which φ_{DC1} performs well. Other k-stage versions of (1) and (2) are straightforward to analyze; however the number of variables specifying such tests prohibits a concise meaningful analysis of their performance.

6. Conclusions

In this paper, we have proposed and analyzed a potentially useful class of multistage tests. It should be noted that, although we have considered only the linear test statistic with normal errors and the sign test statistic, the relative efficiency expressions and numerical data of Sections 3 and 5 are applicable to much broader classes of parametric testing problems, test statistics and error distributions, subject to mild regularity conditions (such as those of Lai (1978), Noether (1955), and/or Paulson (1947)). As demonstrated by the analysis of the above sections, the proposed tests are intermediate in terms of efficiency to fixed-sample-size tests and sequential probability ratio tests. Their implementational complexity and insensitivity to parameter mismatch, however, are comparable to those of fixed-sample-size tests. Thus, the tests proposed here may be preferable to both the fixed-sample-size test and the sequential probability ratio test for many applications.

Acknowledgement

This study was supported by the U.S. National Science Foundation under Grant Number ECS-79-16453.

References

R. Bechhofer (1960). A note on the limiting relative efficiency of the Wald sequential probability ratio test. J. Amer. Statist. Assoc., 55, 660-663.

T. L. Lai (1978). Pitman efficiencies of sequential tests and uniform limit theorems in nonparametric statistics. Ann. Statist., 6, 1027-1047.

G. E. Noether (1955). On a theorem of Pitman. Ann. Math. Statist., 26, 64-68.

E. Paulson (1947). A note on the efficiency of the Wald sequential test. Ann. Math. Statist., 18, 447-450.

A. Wald (1947). Sequential Analysis. New York: Wiley.

θ	$\beta_{FSS}(\theta)$			$n_1 = 8;\ n = 20$					
				$\alpha^* = 0.5$			$\alpha^* = 0.2$		
	M = 10	M = 15	M = 20	$\beta_1(\theta)$	$\beta_2(\theta)$	$E(N\|\theta)$	$\beta_1(\theta)$	$\beta_2(\theta)$	$E(N\|\theta)$
0.0	0.01	0.01	0.01	0.01	0.01	14.00	0.01	0.01	10.40
0.2	0.045	0.060	0.076	0.062	0.076	16.57	0.067	0.074	12.70
0.4	0.144	0.219	0.295	0.220	0.294	18.45	0.244	0.281	15.37
0.6	0.334	0.499	0.639	0.487	0.637	19.46	0.537	0.607	17.65
0.8	0.581	0.780	0.895	0.754	0.892	19.86	0.802	0.862	19.07
1.0	0.798	0.939	0.984	0.919	0.983	19.97	0.943	0.967	19.72

Table 1: Power functions of fixed-sample-size linear tests and double-check linear tests. $\beta_1(\cdot)$ and $\beta_2(\cdot)$ are for double-check Procedures 1 and 2, respectively. The errors are normally distributed and the significance level α equals 0.01.

Table 2: Asymptotic efficiencies under H_0 and H_1 of the double-check tests relative to the fixed-sample-size test. The values in parenthesis are those under H_1.

(a) $\alpha = (1-\beta) = 0.05$

	$\alpha^* = .8\sqrt{\alpha}$		$\alpha^* = \sqrt{\alpha}$		$\alpha^* = 1.5\sqrt{\alpha}$		$\alpha^* = 3\sqrt{\alpha}$	
K	proc #1	proc #2	proc #1	proc #2	proc #1	proc #2	proc #1	proc #2
3.0	1.071 (.427)	1.070 (.427)	1.117 (.485)	1.119 (.486)	1.209 (.628)	1.235 (.642)	1.052 (.799)	1.249 (.949)
2.0	1.202 (.562)	1.209 (.566)	1.257 (.626)	1.279 (.637)	1.284 (.732)	1.390 (.792)	.946 (.741)	1.252 (.980)
1.5	1.255 (.655)	1.283 (.669)	1.287 (.704)	1.355 (.742)	1.232 (.751)	1.429 (.871)	.841 (.676)	1.234 (.991)
1.25	1.259 (.701)	1.316 (.732)	1.266 (.734)	1.381 (.800)	1.163 (.740)	1.428 (.909)	.769 (.628)	1.217 (.995)
1.0	1.227 (.736)	1.335 (.801)	1.200 (.743)	1.388 (.860)	1.050 (.704)	1.406 (.943)	.679 (.575)	1.194 (.998)

(b) $\alpha = (1-\beta) = 10^{-6}$

	$\alpha^* = \sqrt{\alpha}$		$\alpha^* = 1.5\sqrt{\alpha}$		$\alpha^* = 3\sqrt{\alpha}$	
K	proc #1	proc #2	proc #1	proc #2	proc #1	proc #2
2.0	1.466 (.490)	1.467 (.490)	1.511 (.505)	1.512 (.505)	1.597 (.535)	1.597 (.535)
1.5	1.467 (.588)	1.467 (.588)	1.513 (.606)	1.513 (.606)	1.599 (.642)	1.599 (.643)
1.25	1.467 (.653)	1.467 (.653)	1.511 (.673)	1.513 (.674)	1.578 (.704)	1.600 (.714)
1.0	1.417 (.709)	1.468 (.735)	1.405 (.704)	1.514 (.758)	1.352 (.678)	1.601 (.803)
0.8	1.174 (.653)	1.468 (.816)	1.139 (.634)	1.513 (.842)	1.084 (.604)	1.598 (.890)
0.5	.734 (.490)	1.439 (.960)	.712 (.475)	1.461 (.975)	.678 (.453)	1.482 (.990)

Table 2 (continued)

(c) $\alpha = 10^{-6}$, $1-\beta = 0.05$

K	$\alpha^* = \sqrt{\alpha}$		$\alpha^* = 1.5\sqrt{\alpha}$		$\alpha^* = 3\sqrt{\alpha}$	
	proc #1	proc #2	proc #1	proc #2	proc #1	proc #2
3.0	1.820 (.474)	1.821 (.474)	1.916 (.500)	1.915 (.500)	2.102 (.551)	2.103 (.551)
2.0	1.821 (.629)	1.822 (.629)	1.915 (.662)	1.918 (.663)	2.089 (.724)	2.104 (.729)
1.5	1.801 (.743)	1.819 (.750)	1.872 (.771)	1.909 (.786)	1.967 (.808)	2.067 (.849)
1.25	1.747 (.795)	1.805 (.822)	1.784 (.810)	1.883 (.855)	1.806 (.816)	2.004 (.906)

(d) $\alpha = 0.05$, $(1-\beta) = 10^{-6}$

K	$\alpha^* = \sqrt{\alpha}$		$\alpha^* = 1.5\sqrt{\alpha}$		$\alpha^* = 3\sqrt{\alpha}$	
	proc #1	proc #2	proc #1	proc #2	proc #1	proc #2
1.25	1.050 (.597)	1.052 (.598)	1.023 (.645)	1.075 (.678)	.725 (.592)	1.151 (.940)
1.0	1.047 (.641)	1.101 (.673)	.911 (.608)	1.143 (.763)	.638 (.533)	1.175 (.981)
0.8	.912 (.597)	1.142 (.748)	.769 (.542)	1.199 (.845)	.555 (.474)	1.165 (.995)
0.5	.606 (.449)	1.201 (.890)	.522 (.407)	1.239 (.964)	.399 (.355)	1.123 (1.000)

$1-\beta$ \ α	10^{-2}	10^{-3}	10^{-4}	10^{-6}
10^{-2}	.654 (.350)	.580 (.341)	.490 (.343)	.356 (.355)
10^{-3}	.646 (.233)	.598 (.238)	.527 (.249)	.404 (.270)
10^{-4}	.629 (.171)	.611 (.182)	.555 (.196)	.438 (.220)
10^{-5}	.612 (.110)	.630 (.126)	.593 (.140)	.489 (.164)

Table 3: Asymptotic efficiencies of Procedure 2 double-check test relative to the sequential probability ratio test under H_0 and H_1. The values in parentheses are those under H_1. The double-check test has $\alpha^* = 3\sqrt{\alpha}$ and $K = 2$.

p	$\beta_{FSS}(p)$			$n_1 = 8 \quad n = 20$					
				$\alpha^* = 0.4$			$\alpha^* = 0.2$		
	M = 10	M = 15	M = 20	$\beta_1(p)$	$\beta_2(p)$	$E(N\|p)$	$\beta_1(p)$	$\beta_2(p)$	$E(N\|p)$
0.5	0.01	0.01	0.01	0.01	0.01	12.80	0.01	0.01	10.40
0.6	0.043	0.056	0.072	0.062	0.071	15.50	0.064	0.070	12.63
0.7	0.140	0.204	0.287	0.229	0.286	17.89	0.240	0.276	15.40
0.8	0.355	0.512	0.678	0.554	0.676	19.40	0.578	0.652	18.01
0.9	0.707	0.874	0.966	0.894	0.965	19.95	0.910	0.949	19.64

Table 4: Power functions of fixed-sample-size sign tests and double-check sign tests. The value p is the probability of having a positive observation. $\beta_1(\cdot)$ and $\beta_2(\cdot)$ are for double-check Procedures 1 and 2, respectively. The significance level is $\alpha = 0.01$.

Number of stages (k)	$\alpha=1-\beta=10^{-2}$	$\alpha=1-\beta=10^{-4}$	$\alpha=1-\beta=10^{-6}$	$\alpha=10^{-2}$ $1-\beta=10^{-6}$	$\alpha=10^{-6}$ $1-\beta=0.1$
2	1.323 (.730)	1.417 (.716)	1.417 (.709)	1.196 (.658)	1.631 (.838)
3	1.401 (.591)	1.642 (.574)	1.680 (.566)	1.198 (.504)	2.099 (.732)
4	1.386 (.504)	1.748 (.486)	1.847 (.477)	1.143 (.414)	2.442 (.655)
6	1.281 (.396)	1.781 (.378)	1.995 (.369)	1.006 (.310)	2.849 (.551)
8	1.166 (.331)	1.718 (.314)	2.009 (.305)	.886 (.251)	3.023 (.481)
10	1.064 (.287)	1.628 (.270)	1.964 (.262)	.789 (.212)	3.072 (.430)

Table 5: Asymptotic efficiencies of k-stage test relative to the fixed-sample-size test with various error probabilities α and $1-\beta$. The k-stage test has an equal number of samples in each stage and the significance level at each stage is $\alpha^{1/k}$.

END

FILMED

12-85

DTIC